

Multi-Concept Customization of Text-to-Image Diffusion

Year: 2022 | Citations: 1117 | Authors: Nupur Kumari, Bingliang Zhang, Richard Zhang, Eli Shechtman, Jun-Yan Zhu

Abstract

While generative models produce high-quality images of concepts learned from a large-scale database, a user often wishes to synthesize instantiations of their own concepts (for example, their family, pets, or items). Can we teach a model to quickly acquire a new concept, given a few examples? Furthermore, can we compose multiple new concepts together? We propose Custom Diffusion, an efficient method for augmenting existing text-to-image models. We find that only optimizing a few parameters in the text-to-image conditioning mechanism is sufficiently powerful to represent new concepts while enabling fast tuning (~ 6 minutes). Additionally, we can jointly train for multiple concepts or combine multiple fine-tuned models into one via closed-form constrained optimization. Our fine-tuned model generates variations of multiple new concepts and seamlessly composes them with existing concepts in novel settings. Our method outperforms or performs on par with several baselines and concurrent works in both qualitative and quantitative evaluations, while being memory and computationally efficient.